JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS

ELSEVIER Journal of Computational and Applied Mathematics 75 (1996) 1–21

Extension of Henrici's method to matrix sequences

M.D. Benchiboun *

Laboratoire d'Analyse Numérique et d'Optimisation, Université des Sciences et Technologies de Lille, 59655 Villeneuve d'Ascq cedex, France

Received 2 February 1995; revised 19 February 1996

Abstract

In this paper we generalize the definition of linear convergence to matrix sequences. This new definition is used to establish some new results useful to study the new extension of Henrici's method. A convergence theorem, an algorithm for implementation of this method and some numerical examples are given.

Keywords: Henrici's method; Linear convergence; Matrix sequences

AMS classification: 65B05, 65F30

0. Introduction and notations

As in the scalar and the vector cases, the matrix sequences occur most naturally in numerical analysis:

- In the digital sum of power series in a matrix. In this case, the suggestion of Golub and Van Loan [8] for computing a function $f(A)$ of a $p \times p$ matrix is to approximate it with the sequence of matrices $g_k(A)$, where g_k might be a truncated Taylor series approximate of order k to f. A source of such problems lies in the theory of differential equation of type $X''(t) = -AX(t)$ where A and $X(t)$ are $p \times p$ matrices.
- To solve the linear matrix equtions $X - AXD = C$ (which are studied in more general situations by Wimmer [16]), we construct S_n as $S_{n+1} = AS_nD + C$, which converges, under some conditions, to the solution of the equation.
- To calculate the value of continued fraction of matrix. In [6], Busby and Fair study a continued fraction

$$\frac{I}{I+}\frac{A}{I+}\frac{A}{I+\ldots}$$

* Fax: 3320434173; e-mail: benchib@iut-a.univ-lillel.fr.

SSDI 0377-0427(96)00031-3

where A is an element of a banach algebra. This continued fraction converges , under some assumptions to an element B where AB is a root of the polynomial $X^2 + X - A$. If A is a $p \times p$ matrix it is possible to construct a sequence of matrices A_k (by the convegent of order k of the continued fraction), which converges to a matrix B where AB is the root of the polynomial matrix $X^2 + X - A$.

Matrix sequences play a central role in numerical analysis. It converge, under certain assumptions, to the solution of the problem but, in some cases, the convergence is slow. So, it is natural to extend the extrapolation methods to the matricial case. Wynn, in [18], suggests (as far as we know) the only extrapolation method applied to matrices, which is a matricial version of the epsilon algorithm. This method was applied by Brezinski [2] to compute the inverse of matrices and to accelerate the convergence of some matrix sequences, and by Tan and Andrew [14] for numerical computation of partial derivatives of several eigenvalues and eigenvectors of a matrix which depends on a number of parameters. The disadvantage of this method is that it requires to compute, at each step, the inverse or the pseudo-inverse of a matrix. It is explanable that Wynn uses the vectorial epsilon algorithm instead of the matrix one avoiding the computation of the inverse of a matrix at each step of the algorithm.

For this reason, it is fundamental to construct other extrapolation methods for accelerating the convergence of the matrix sequences (S_n). Such a construction needs sufficient information about (S_n) and the error sequences $(S_n - S)$, where S is the limit of (S_n). In the scalar and vector cases, two important types of convergent sequence were defined: linearly and logarithmically convergent sequences. Such sequences were studied by several authors.

In a recent paper by the present author [1], a study and applications of linear convergence of vector sequences and relations between the definitions proposed in [10, 11, 13] were given. We now extend this work in two ways. Firstly, we generalize the definition of linear convergence for matrix sequences, and we give some results about this convergence. Using this definition, we generalize a theorem given in [1, 7, 17], which gives the relation between the asymptotic behaviour of the ratios of the errors $(S_n - S)$ and that of the differences $(S_{n+1} - S_n)$ for linearly convergent scalar sequences. Such a result is important because it allows us to recognize the occurrence of the linear convergence in matrix sequences. We also give some properties about the matrix B used in the definition of linear convergence of matrix sequences, which is not unique in the general case. This remark obviously merits explanation.

Secondly, we give for the first time as far as we know, an extension of Henrici's transformation [9] to matrix sequences. One of the reason of our interest in this method is the paper of Sadok [13], where he made a comparison study of the vector epsilon algorithm, topological epsilon algorithm and Henrici's method for vector sequences, and he showed by examples that the Henrici's method is more efficient than the others. The advantage of the Henrici's method to matrix sequences is the easier way to implement it with the H-algorithm, it involves only addition of matrices and multiplication by scalars.

Henrici's method to matrix sequences is defined as a ratio of determinants. Applying the extension of the Schur complement [3] we obtain a new formula which is used to prove a convergence acceleration theorem. We also give a way for implementing this transformation based on the H-algorithm [4, 5] and some numerical examples. These examples consist in calculating the partial sum of a power series in a matrix, the solution of the matrix equation $X - AXD = C$, and the inversion of $I - D$ for every nonsingular matrix D such that $\| D \|$ is less than 1.

In this paper we use the following notation. We denote by $U=(u_1,u_2,\ldots,u_p)^{\mathrm{T}} \in \mathbb{C}^p$ a complex vector, by $A=(a_{i,j})_{i,j} \in M_{p\times q}(\mathbb{C})$ a complex matrix, and by $B=(b_{i,j})_{i,j} \in M_{p\times p}(\mathbb{C})$ a complex square matrix. We also use the following vector norms: $\| U \|_\infty = \max_{1\leqslant i\leqslant p} |u_i|$, $\| U \|_1 = \sum_{i=1}^{p} |u_i|$, $\| U \|_2 = (\sum_{i=1}^{p} | u_i |^2)^{1/2}$. We let $\| B \|_2$ be the induced matrix norm, and we consider the following matrix norms: $\| A \|_1 = \mathrm{Max}_{1\leqslant j\leqslant q} \sum_{i=1}^{p} | a_{i,j} |$, and $\| A \|_{\mathrm{F}} = (\sum_{i=1,j=1}^{i=p,j=q} | a_{i,j} |^2)^{1/2}$ (the Frobenius matrix norm), and we shall use the natation $\| \,.\, \|$, where $\| \,.\, \| = \| \,.\, \|_1$ or $\| \,.\, \|_{\mathrm{F}}$.

We define $(S_n) = (s_{ij}(n))_{ij} \in M_{p\times q}(\mathbb{C})$ to be a sequence of complex matrices converging to S, and we denote by $(E_n) = (S_n - S) = (e_{ij}(n))_{ij}$ and $(\Delta S_n) = (S_{n+1} - S_n) = (\Delta s_{ij}(n))_{ij}$ respectively the sequence of the errors and the differences.

1. Linear convergence to matrix sequences

This section has two objects. The first is to define linear convergence for matrix sequences and its properties. The second is to give some results which are important for the next section.

1.1. Definitions and examples

Recall that a sequence (s_n) of numbers converges linearly if
- there exists a number s such that $\lim_{n\to\infty} s_n = s$,
- there exists $N \in \mathbb{N}$ such that $\forall n \geqslant N$, $s_n \neq s$ and,
- there exists a number r such that $-1 \leqslant r < 1$ and

$$\lim_{n\to\infty} \frac{s_{n+1}-s}{s_n - s} = r.$$

Let us now consider a sequence of complex matrices $(S_n) \in M_{p\times q}(\mathbb{C})$. By analogy with the vector case [1, 10], we propose the following definition of linear convergence.

Definition 1.1. Let $(S_n) \in M_{p\times q}(\mathbb{C})$ be a sequence converging to S. The convergence of (S_n) is linear if and only if
- there exists $N \in \mathbb{N}$ such that $\forall n \geqslant N$ $S_n \neq S$ and,
- there exists a nonsingular matrix $B \in M_{p\times p}(\mathbb{C})$ with $\| B \| < 1$ such that

$$\lim_{n\to\infty} \frac{(S_{n+1}-S) - B(S_n - S)}{\| S_n - S \|} = 0. \tag{1}$$

For each nonsingular matrix B such that $\| B \|$ is less than 1 we call L_B the set of convergent sequences which satisfy (1).

We remark that in some cases the matrix B is not unique.

Example 1.2. Let (S_n) be the matrix sequence defined by

$$S_n = \begin{pmatrix} (\frac{1}{2})^n & (\frac{1}{3})^n \\ (\frac{1}{4})^n & (\frac{1}{5})^n \end{pmatrix},$$

then (S_n) converges to 0 and, using $\| \cdot \|_F$, it follows that $(S_n) \in L_B$, where

$$B = \begin{pmatrix} \frac{1}{2} & c \\ 0 & d \end{pmatrix},$$

for all $c \in \mathbb{C}$ and $d \in \mathbb{C} - \{0,1\}$ such that $c^2 + d^2 < \frac{3}{4}$.

Example 1.3. Let us now consider the matrix A such that $\| A \|$ is less than 1, then $I - A$ is nonsingular and $(I - A)^{-1} = \sum_{i=0}^{\infty} A^i$ is the limit of $S_n = \sum_{i=0}^{n} A^i$. In this case, if A is nonsingular we have $(S_n) \in L_A$.

1.2. Some results about linearly convergent sequences

Let us first propose the following theorem, which gives the asymptotic behaviour of the ratios $\| E_{n+1} \| / \| E_n \|$ and $\| \Delta S_n \| / \| E_n \|$. This theorem will be used for proving other results.

Theorem 1.4. *Let $(S_n) \in M_{p\times q}(\mathbb{C})$ be a sequence converging to S. If there exists a nonsingular matrix $B \in M_{p\times p}(\mathbb{C})$ such that $\| B \|$ is less than 1 and $(S_n) \in L_B$ (S_n is linearly convergent), then the following inequalities hold:*

$$\frac{1}{\| B^{-1} \|} \leqslant \liminf_{n\to\infty} \frac{\| E_{n+1} \|}{\| E_n \|} \leqslant \limsup_{n\to\infty} \frac{\| E_{n+1} \|}{\| E_n \|} \leqslant \| B \| \tag{2}$$

and

$$1 - \frac{1}{\| B^{-1} \|} \leqslant \liminf_{n\to\infty} \frac{\| \Delta S_n \|}{\| E_n \|} \leqslant \limsup_{n\to\infty} \frac{\| \Delta S_n \|}{\| E_n \|} \leqslant \| B \| + 1. \tag{3}$$

Proof. Let $(S_n) \in L_B$. There exists a matrix sequence $\alpha_n = (\alpha_{i,j}(n))_{i,j}$ converging to 0 such that

$$E_{n+1} - BE_n = \| E_n \| \alpha_n. \tag{4}$$

For $\varepsilon > 0$ there exists an integer N such that

$$\| \alpha_n \| < \varepsilon \quad \text{for all } n > N. \tag{5}$$

From (4) we have $E_{n+1} = \| E_n \| \alpha_n + BE_n$ and the triangle inequality gives

$$\left| \frac{\| BE_n \|}{\| E_n \|} - \| \alpha_n \| \right| \leqslant \frac{\| E_{n+1} \|}{\| E_n \|} \leqslant \| B \| + \| \alpha_n \|. \tag{6}$$

We also have

$$\frac{\| BE_n \|}{\| E_n \|} = \frac{\| BE_n \|}{\| B^{-1}BE_n \|} \geqslant \frac{1}{\| B^{-1} \|}. \tag{7}$$

From (5)–(7), and choosing $\varepsilon < 1/\| B^{-1} \|$ there exists N_ε such that

$$\left| \frac{1}{\| B^{-1} \|} - \varepsilon \right| \leqslant \frac{\| E_{n+1} \|}{\| E_n \|} \leqslant \| B \| + \varepsilon \quad \text{for all } n > N_\varepsilon,$$

which ends the proof of (2).

Using (2) and $\Delta S_n = E_{n+1} - E_n$, (3) follows. □

Remark 1.5. (a) Given $(S_n) \in L_B$ with B nonsingular and $\|B\|$ less than 1. If $\lim_{n\to\infty} \|E_{n+1}\|/\|E_n\| = R$ exists then, from Theorem 1.4, we obtain $0 < 1/\|B^{-1}\| \leqslant R \leqslant \|B\| < 1$.

In general, the converse is not true, as proved by the following example. Let (S_n) be a sequence of matrices defined by

$$S_n = \begin{pmatrix} (\frac{1}{2})^n \cos(\frac{1}{2}n\pi) & (\frac{1}{2})^n \\ (\frac{1}{3})^n & (\frac{1}{4})^n \end{pmatrix}.$$

Then (S_n) converges to 0, and if we use the Frobenius matrix norm $\| \cdot \|_F$, we may prove that $\lim_{n\to\infty} \|E_{n+1}\|/\|E_n\| = \frac{1}{2}$, but there is no matrix B such that $(S_n) \in L_B$.

(b) If there exist two matrices B and B_n, with B nonsingular and $\|B\|$ less than 1, and B_n converging to 0 such that $E_{n+1} = (B + B_n)E_n$, then

$$\lim_{n\to\infty} \frac{E_{n+1} - BE_n}{\|E_n\|} = 0,$$

and hence $(S_n) \in L_B$. The converse is not always true: from Example 1.2 we know that $S_n \in L_B$ but we cannot find two matrices B' and B'_n, with B' nonsingular and $\|B'\| < 1$, and B'_n convergent to 0 such that $E_{n+1} = (B' + B'_n)E_n$.

1.3. Relations between the sequences (E_n) and (ΔS_n)

The definition of linear convergence (1) depends on the sequence of errors ($E_n = S_n - S$). Therefore, if we have not some information about (E_n), we cannot know whether (S_n) converges linearly or not. That is the reason why, in this section, we propose a new definition, obtained from (1) by substituting (ΔS_n) for (E_n). This definition is equivalent to the first one, as we prove in Theorem 1.9, and will be used in the sequel.

Let us first recall the relations between (E_n) and (ΔS_n) in the scalar and vector cases.

(a) *Scalar case*. Let $(s_n) \in \mathbb{C}$ be a scalar sequence converging to s, then we have the following result:

Theorem 1.6 (Delahaye [7] and Wimp [17]). *Given $r \in \mathbb{C}$, $|r| < 1$, then*

$$\lim_{n\to\infty} \frac{e_{n+1}}{e_n} = r \quad \textit{if and only if} \quad \lim_{n\to\infty} \frac{\Delta s_{n+1}}{\Delta s_n} = r.$$

(b) *Vector case.*

Theorem 1.7 (Benchiboun [1]). *Let $(s_n) \in \mathbb{C}^p$ be a vector sequence converging to $s \in \mathbb{C}^p$. Given $B \in M_{p\times p}(\mathbb{C})$ nonsingular and $\|B\| < 1$, then*

$$\lim_{n\to\infty} \frac{e_{n+1} - Be_n}{\|e_n\|} = 0 \quad \textit{if and only if} \quad \lim_{n\to\infty} \frac{\Delta s_{n+1} - B\Delta s_n}{\|\Delta s_n\|} = 0.$$

(c) *Matrix case.*
Let us now consider a sequence $(S_n) \in M_{p\times q}(\mathbb{C})$ converging to $S \in M_{p\times q}(\mathbb{C})$. Define

$$D_n = \frac{\Delta S_{n+1}}{\| \Delta S_n \|} \in M_{p\times q}(\mathbb{C}),$$

and $C_n = \sum_{i=0}^{\infty} A_{n,i}$ be a series with $A_{n,0} = D_n$ and $A_{n,i} = \| D_n \| \, \| D_{n+1} \| \cdots \| D_{n+i-1} \| \, D_{n+i}$ for $i \geqslant 1$.
We first give the following lemma.

Lemma 1.8. *If there exists $B \in M_{p\times p}(\mathbb{C})$ nonsingular and $\| B \| < 1$ such that*

$$\lim_{n\to\infty} \frac{\Delta S_{n+1} - B\Delta S_n}{\| \Delta S_n \|} = 0$$

then,

(i) $$\frac{1}{\| B^{-1} \|} \leqslant \liminf_{n\to\infty} \frac{\| \Delta S_{n+1} \|}{\| \Delta S_n \|} \leqslant \limsup_{n\to\infty} \frac{\| \Delta S_{n+1} \|}{\| \Delta S_n \|} \leqslant \| B \|, \tag{8}$$

(ii) *the series C_n converges normally,*

(iii) $$\frac{1}{\| B^{-1} \| \, \| B - I \|} \leqslant \liminf_{n\to\infty} \| C_n \| \leqslant \limsup_{n\to\infty} \| C_n \| \leqslant \frac{\| B \|}{1 - \| B \|}. \tag{9}$$

Proof. The assertion (i) is obvious, and (ii) follows from (i).
(iii): From (i), it can be shown that

$$\limsup_{n\to\infty} \| C_n \| \leqslant \frac{\| B \|}{1 - \| B \|}.$$

We also have

$$\| (B - I)C_n \| \geqslant \left| \, \| (B - I)C_n + B\frac{\Delta S_n}{\| \Delta S_n \|} \| - \| B\frac{\Delta S_n}{\| \Delta S_n \|} \| \, \right|,$$

and we can prove that

$$\lim_{n\to\infty} \left((B - I)C_n + B\frac{\Delta S_n}{\| \Delta S_n \|}) \right) = 0. \tag{10}$$

So, for all ε such that $0 < \varepsilon < 1/\| B^{-1} \|$, there exists N such that, for $n > N$,

$$\| (B - I)C_n \| > \| B\frac{\Delta S_n}{\| \Delta S_n \|} \| - \varepsilon = \frac{\| B\Delta S_n \|}{\| B^{-1}B\Delta S_n \|} - \varepsilon,$$

therefore, for such an index,

$$\| (B - I) \| \, \| C_n \| > \frac{1}{\| B^{-1} \|} - \varepsilon,$$

and finally we obtain

$$\lim_{n\to\infty} \inf \| C_n \| \geqslant \frac{1}{\| B^{-1} \|} \frac{1}{\| (B-I) \|}. \qquad \square$$

We use this lemma for proving the following theorem.

Theorem 1.9. *Let $S_n \in M_{p\times q}(\mathbb{C})$ be a sequence converging to S and $B \in M_{p\times p}(\mathbb{C})$ be a nonsingular matrix such that $\| B \|$ less than* 1, *then*

$$\lim_{n\to\infty} \frac{E_{n+1} - BE_n}{\| E_n \|} = 0 \quad \textit{if and only if} \quad \lim_{n\to\infty} \frac{\Delta S_{n+1} - B\Delta S_n}{\| \Delta S_n \|} = 0.$$

Proof. ($\Rightarrow$): The result follows from (2) and (3) of Theorem 1.4.
($\Leftarrow$): We have

$$S_{n+j+2} - S_{n+j+1} = \| \Delta S_n \| \, \| D_n \| \ldots \| D_{n+j-1} \| D_{n+j}, \quad \text{for } j = 0, 1, 2, \ldots$$

By summation we obtain

$$S_{n+i+2} - S_{n+1} = \| \Delta S_n \| \sum_{j=0}^{i} A_{n,j} \quad \text{for all } i.$$

From Lemma 1.8, the series C_n converges, therefore, when i tends to infinity, it follows that

$$E_{n+1} = - \| \Delta S_n \| C_n$$

and thus

$$E_{n+1} - BE_n = \| \Delta S_n \| \left((B-I)C_n + B\frac{\Delta S_n}{\| \Delta S_n \|} \right).$$

Finally, from (3) and (10) we obtain

$$\lim_{n\to\infty} \frac{E_{n+1} - BE_n}{\| E_n \|} = 0 \qquad \square$$

Remark 1.10. Theorem 1.9 is important because it allows us to recognize, through the sequence (ΔS_n), whether or not (S_n) converges linearly.

For each nonsingular matrix B such that $\| B \|$ is less than 1, we obtain, from Theorem 1.9,

$$L_B = \left\{ (S_n) \text{ converging to } S \text{ such that } \lim_{n\to\infty} \frac{\Delta S_{n+1} - B\Delta S_n}{\| \Delta S_n \|} = 0 \right\}.$$

Very often, it is easy to verify if (S_n) belongs to L_B or not.

1.4. Properties of the matrix B

In the general case the matrix B is not unique (see Example 1.2).

In order to determine the relations between the two matrices B and B' such that (S_n) belongs to $L_B \bigcap L_{B'}$ and the conditions under which B is unique, we first give the following theorem:

Theorem 1.11. *Let (S_n) be a sequence converging to S. If there exists a matrix $B \in M_{p\times p}(\mathbb{C})$ such that $(S_n) \in L_B$, then for every matrix $B' \in M_{p\times p}(\mathbb{C})$*

$$(S_n) \in L_{B'} \quad \textit{if and only if} \quad \lim_{n\to\infty} \frac{(B-B')\Delta S_n}{\| \Delta S_n \|} = 0. \tag{11}$$

1.4.1. A sufficient condition for uniqueness of the matrix B

In this section we consider the square matrix sequence $(S_n) \in M_{p\times p}(\mathbb{C})$. Let us first give an example for which B is unique.

Example 1.12. Let (S_n) be a sequence defined by

$$S_n = \begin{pmatrix} (\frac{1}{2})^n & (\frac{1}{3})^n \\ (\frac{1}{3})^n & (\frac{1}{2})^n \end{pmatrix}.$$

(S_n) converges to 0, and using the Frobenius matrix norm $\| . \|_F$ we may prove that $B = \mathrm{diag}(\frac{1}{2}, \frac{1}{2})$ is the unique matrix such that $(S_n) \in L_B$.

Let us now study a sufficient condition that assures the uniqueness of B. Before we do this, the following lemma is needed.

Lemma 1.13 (Ostrowski [12, p. 168]). *Let $A \in M_{p\times p}(\mathbb{C})$ be a matrix. Then*

$$\| A^{-1} \|_2 \leqslant \frac{\| A \|_F^{p-1}}{| \det(A) | \, (p-1)^{(p-1)/2}}.$$

This lemma will be used to prove the following theorem.

Theorem 1.14. *Given $(S_n) \in M_{p\times p}(\mathbb{C})$, let $B \in M_{p\times p}(\mathbb{C})$ be a nonsingular matrix with $\| B \| < 1$, such that $(S_n) \in L_B$. If there exists a real number $d > 0$ and an integer $N > 0$ such that $| \det(\Delta S_n) | \geqslant d \, \| \Delta S_n \|^p$, for $n > N$, then the matrix B is unique.*

Proof. Suppose that there exists a nonsingular matrix B' with $\| B' \| < 1$, such that $S_n \in L_{B'}$, then, from Theorem 1.11, there exists a matrix sequence (ε_n) converging to 0 such that

$$\frac{(B-B')\Delta S_n}{\| \Delta S_n \|} = \varepsilon_n. \tag{12}$$

By assumption the matrix $\Delta S_n / \| \Delta S_n \|$ is nonsingular for n larger than N.

Therefore, from (12) and using the induced matrix norm $\| \,.\, \|_2$ and Lemma 1.13, we can show that

$$\| B - B' \|_2 \leqslant \| \varepsilon_n \|_2 \frac{1}{d(p-1)^{(p-1)/2}} \quad \text{for } n > N.$$

So, when n tends to infinity, we obtain $B = B'$. □

Using the notation $(S_n) = (s_{ij}(n))_{i,j}$ and $(\Delta S_n) = (\Delta s_{ij}(n))_{i,j}$, we have the following corollary.

Corollary 1.15. *Given $(S_n) \in M_{p\times p}(\mathbb{C})$, let $B \in M_{p\times p}(\mathbb{C})$ be a nonsingular matrix with $\| B' \| < 1$, such that $(S_n) \in L_B$. If $\lim_{n\to\infty} \Delta s_{ij}(n)/\| \Delta S_n \| = c_{ij}$ exists for all $i,j = 1,2,\ldots,p$, and if the matrix $C = (c_{ij})_{i,j}$ is nonsingular, then the matrix B is unique.*

Proof. The result follows from Theorem 1.14. □

For Example 1.12, using the Frobenius matrix norm, the matrix $C = \text{diag}(-\frac{1}{2}\sqrt{2}, -\frac{1}{2}\sqrt{2})$ is nonsingular and B is unique.

Example 1.16. Let (S_n) be the sequence defined by

$$S_n = \begin{pmatrix} (-\frac{1}{2})^n & (\frac{1}{3})^n \\ (\frac{1}{3})^n & (\frac{1}{2})^n \end{pmatrix}.$$

Then (S_n) converges to 0 and, using the Frobenius matrix norm $\| \,.\, \|_F$, we can prove that $B = \text{diag}(-\frac{1}{2}, \frac{1}{2})$ is the unique matrix such that $(S_n) \in L_B$.

Note that the sequence (S_n) satisfies the conditions of Theorem 1.14 but not those of Corollary 1.15 because $\Delta s_{11}(n)/\| \Delta S_n \|$ does not converge.

1.4.2. The diagonal case

In this section, we consider the sequences $(S_n) \in M_{p\times q}(\mathbb{C})$ for which B is a diagonal matrix.
We have the following result.

Proposition 1.17. *Let $(S_n) \in M_{p\times q}(\mathbb{C})$ be a sequence converging to S such that $\lim_{n\to\infty} \Delta s_{ij}(n)/\| \Delta S_n \| = c_{ij}$ exists for all $i = 1,2,\ldots,p$ and $j = 1,2,\ldots,q$. Let $B = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_p)$ and $B' = \text{diag}(\lambda_1', \lambda_2', \ldots, \lambda_p')$, two diagonal matrices such that $S_n \in L_B \bigcap L_{B'}$. If there exist i,j such that $c_{ij} \neq 0$, then the diagonal matrices B and B' coincide.*

Proof. The result follows from Theorem 1.11. □

Remark 1.18. From Proposition 1.17 we remark that, if $(S_n) \in L_B$ with B a diagonal matrix and, if for all $i \in \{1,2,\ldots,p\}$ there exists $j \in \{1,2,\ldots,q\}$ such that $c_{ij} \neq 0$ then, B is unique.

2. Matrix version of Henrici's transformation

The aim of this section is to define and to study the matrix version of Henrici's transformation. This method is implemented and applied to some matrix sequences.

2.1. Definitions

We shall first recall the definition of Henrici's transformation in the vector case. For a vector sequence $(s_n) \in \mathbb{C}^p$ Henrici proposed the following transformation [9, formula 5-3, p. 116] which is studied in [13]:

$$h:(s_n) \to (h_n); \qquad h_n = s_n - X_n(\Delta X_n)^{-1}\Delta s_n,$$

where X_n is the matrix whose columns are $\Delta s_n, \ldots, \Delta s_{n+p-1}$.

Sadok [13, formula (2), p. 103] noticed that h_n can be expressed as a ratio of two determinants.

Let $(S_n) \in M_{p\times q}(\mathbb{C})$ be a matrix sequence converging to S and $U = (u_1, u_2, \ldots, u_q)^{\mathrm{T}} \in \mathbb{C}^q$ be a complex vector. We consider the following matrix sequence transformation $H:(S_n) \to (H_n)$ defined by the ratio of two determinants

$$H_n = \begin{vmatrix} S_n & \cdots & S_{n+p} \\ \sum_{j=1}^{q} \Delta s_{1j}(n)u_j & \cdots & \sum_{j=1}^{q} \Delta s_{1j}(n+p)u_j \\ \vdots & & \vdots \\ \sum_{j=1}^{q} \Delta s_{pj}(n)u_j & \cdots & \sum_{j=1}^{q} \Delta s_{pj}(n+p)u_j \end{vmatrix} \div \begin{vmatrix} 1 & \cdots & 1 \\ \sum_{j=1}^{q} \Delta s_{1j}(n)u_j & \cdots & \sum_{j=1}^{q} \Delta s_{1j}(n+p)u_j \\ \vdots & & \vdots \\ \sum_{j=1}^{q} \Delta s_{pj}(n)u_j & \cdots & \sum_{j=1}^{q} \Delta s_{pj}(n+p)u_j \end{vmatrix}, \tag{13}$$

where the generalized determinant in the numerator denotes the matrix obtained by expanding it with respect to its first row, (that is a combination of the matrices $S_n, S_{n+1}, \ldots, S_{n+p}$).

Then H_n can also be expressed as follows:

$$H_n = \begin{vmatrix} S_n & \Delta S_n & \cdots & \Delta S_{n+p-1} \\ \Delta S_n U & \Delta^2 S_n U & \cdots & \Delta^2 S_{n+p-1} U \end{vmatrix} \div \left| \Delta^2 S_n U \quad \cdots \quad \Delta^2 S_{n+p-1} U \right|, \tag{14}$$

where $\Delta S_{n+i}U$ is the vector obtained by multiplying the matrix ΔS_{n+i} by the vector U, and $\Delta^2 S_n = \Delta S_{n+1} - \Delta S_n$. Of course the nonsingularity of the matrix $(\Delta^2 S_n U, \Delta^2 S_{n+1} U, \Delta^2 S_{n+p-1} U)$ is assumed, for all n.

Denote by S_n^l, H_n^l the lth columns respectively of S_n and H_n, for $l = 1, 2, \ldots, q$. Therefore, from (14), we obtain

$$H_n^l = \begin{vmatrix} S_n^l & \Delta S_n^l & \cdots & \Delta S_{n+p-1}^l \\ \Delta S_n U & \Delta^2 S_n U & \cdots & \Delta^2 S_{n+p-1} U \end{vmatrix} \div \left| \Delta^2 S_n U \quad \cdots \quad \Delta^2 S_{n+p-1} U \right|. \tag{15}$$

Applying the extension of the Schur complement for the vector case [3] to H_n^l, we obtain

$$H_n^l = S_n^l - (\Delta S_n^l, \Delta S_{n+1}^l, \dots, \Delta S_{n+p-1}^l)^T \star (\Delta^2 S_n U, \Delta^2 S_{n+1} U, \dots, \Delta^2 S_{n+p-1} U)^{-1} \Delta S_n U \tag{16}$$

where $X \star a = a \star X$ stands for the linear combination $a_1X_1 + a_2X_2 + \cdots + a_pX_p$ if $a = (a_1, a_2, \dots, a_p)^{\mathrm{T}} \in \mathbb{C}^p$ and $X = (X_1, X_2, \dots, X_p)^{\mathrm{T}}$, where $X_i \in \mathbb{C}^p$.

Let us use the same definition in the matrix case, that is, $D \star a = a \star D = a_1D_1 + a_2D_2 + \cdots + a_pD_p$, if $a = (a_1, a_2, \dots, a_p)^{\mathrm{T}} \in \mathbb{C}^p$ and $D = (D_1, D_2, \dots, D_p)^T$ where $D_i \in M_{p\times q}(\mathbb{C})$.

Thus, from (16), we immediately obtain

$$H_n = S_n - (\Delta S_n, \Delta S_{n+1}, \dots, \Delta S_{n+p-1})^{\mathrm{T}} \star (\Delta^2 S_n U, \Delta^2 S_{n+1} U, \dots, \Delta^2 S_{n+p-1} U)^{-1} \Delta S_n U. \tag{17}$$

If we set $q = 1$ and $U = 1$, then H is Henrici's transformation in the vector case [9].

2.2. Study of the matrix version of Henrici's transformation

In this section we shall give some results about the transformation H when applied to matrix sequences which are linearly convergent.

Let (S_n) be in L_B with B nonsingular and $\| B \| < 1$, and let $U \in \mathbb{C}^q$. We shall use the following notation: (V_n) a sequence of vectors defined by $V_n = \Delta S_n U$ and (A_n) a sequence of matrices defined by $A_n = (V_n, BV_n, \dots, B^{p-1}V_n)$.

In order to prove a convergence acceleration theorem of the Henrici's transformation in the matrix case, we first give the following result:

Lemma 2.1. *Given $(S_n) \in L_B$ such that $B \in M_{p\times p}(\mathbb{C})$ is nonsingular and $\| B \| < 1$. If there exists $N > 0$ such that the matrix A_n is nonsingular for all $n > N$, then*

$$(I, B, \dots, B^{p-1})^T \star ((A_n)^{-1}(I - B)^{-1} V_n) = (I - B)^{-1} \quad \textit{for } n > N. \tag{18}$$

Proof. Let $Q(t) = (-1)^p(t^p - q_1t^{p-1} - \cdots - q_p)$ be the characteristic polynomial of B. We know, from the Cayley–Hamilton theorem, that

$$B^p = q_1B^{p-1} + q_2B^{p-2} + \cdots + q_pI \tag{19}$$

So

$$B^pV_n = q_1B^{p-1}V_n + q_2B^{p-2}V_n + \cdots + q_pV_n. \tag{20}$$

By assumption the vectors $V_n, BV_n, \dots, B^{p-1}V_n$ are linearly independent for $n > N$. Let us now denote by $a_i^{(0)}, a_i^{(1)}, \dots, a_i^{(p-1)}$ the coefficient of the vector $B^{p+i}V_n$ in the basis $V_n, BV_n, \dots, B^{p-1}V_n$.

From (19) and (20) we obtain

$$(a_{i+1}^{(0)}, a_{i+1}^{(1)}, \dots, a_{i+1}^{(p-1)})^{\mathrm{T}} = Y(a_i^{(0)}, a_i^{(1)}, \dots, a_i^{(p-1)})^{\mathrm{T}} \quad \textit{for } i \geqslant 0,$$

where $(a_0^{(0)}, a_0^{(1)}, \dots, a_0^{(p-1)}) = (q_p, q_{p-1}, \dots, q_1)$ and Y belonging to $M_{p\times p}(\mathbb{C})$ is the matrix defined by

$$Y = \begin{pmatrix} 0 & 0 & \cdots & 0 & q_p \\ 1 & 0 & \cdots & 0 & q_{p-1} \\ 0 & 1 & \cdots & 0 & q_{p-2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & q_1 \end{pmatrix}. \tag{21}$$

It follows that

$$(a_i^0, a_i^1, \dots, a_i^{p-1})^{\mathrm{T}} = Y^i (q_p, q_{p-1}, \dots, q_1)^{\mathrm{T}}. \tag{22}$$

Let d_0, $d_1, \dots$, d_{p-1} be the coefficients of $(I-B)^{-1}V_n$ in the basis V_n, BV_n, $\dots$, $B^{p-1}V_n$. We have

$$(I-B)^{-1}V_n = d_0 V_n + d_1 B V_n + \cdots + d_{p-1} B^{p-1} V_n, \tag{23}$$

and, since $\| B \| < 1$ then, $(I-B)^{-1}V_n = \sum_{i=0}^{\infty} B^i V_n$.

From (22) and (23) we can prove that

$$(q_p, q_{p-1}, \dots, q_1)^{\mathrm{T}} = (I - Y)(d_0 - 1, d_1 - 1, \dots, d_{p-1} - 1)^{\mathrm{T}}. \tag{24}$$

Using (19) and (24) we obtain

$$d_{p-1}B^{p-1} + d_{p-2}B^{p-2} + \cdots + d_1 B + d_0 I = (I-B)^{-1},$$

that is

$$(I, B, \dots, B^{p-1})^{\mathrm{T}} \star (d_0, d_1, \dots, d_{p-1})^{\mathrm{T}} = (I-B)^{-1}. \tag{25}$$

Finally, from (23) and (25), the result follows. □

We also have the following lemma.

Lemma 2.2 (Wilkinson [15]). *Let $X, Y \in M_{p\times p}$ be two square matrices. Assume that X is nonsingular and $\| X^{-1} \| \, \| X - Y \| < 1$, then Y is nonsingular and*

$$\| X^{-1} - Y^{-1} \| \leqslant \frac{\| X^{-1} \|^2 \| X - Y \|}{1 - \| X^{-1} \| \, \| X - Y \|}.$$

Let $(S_n) \in L_B$, and set

$$D_n = \frac{1}{\| \Delta S_n \|_{\mathrm{F}}} A_n = \left(\frac{\Delta S_n}{\| \Delta S_n \|_{\mathrm{F}}} U, \frac{B \Delta S_n}{\| \Delta S_n \|_{\mathrm{F}}} U, \dots, \frac{B^{p-1} \Delta S_n}{\| \Delta S_n \|_{\mathrm{F}}} U \right) \quad \text{and} \quad \Theta_n = \frac{\Delta S_{n+1} - B \Delta S_n}{\| \Delta S_n \|_{\mathrm{F}}}.$$

For $(S_n) \in L_B$, the matrix sequence (Θ_n) converges to 0.

We shall use Lemmas 2.1 and 2.2, and $\|\,.\,\|_F$ to prove the following theorem.

Theorem 2.3. *Let $(S_n) \in L_B$ with B nonsingular and $\| B \|_F < 1$. If there exists $N > 0$ such that D_n is nonsingular for $n > N$ and, if $\lim_{n\to\infty} \| D_n^{-1} \|_F^2 \| \Theta_{n+i} \|_F = 0$, for $i = 0, 1, \ldots, p-1$, then*

$$\lim_{n\to\infty} \frac{\| H_n - S \|_F}{\| S_n - S \|_F} = 0.$$

Proof. Let $(S_n) \in L_B$. From (17) we have

$$H_n - S = (S_n - S) - (\Delta S_n, \Delta S_{n+1}, \ldots, \Delta S_{n+p-1})^{\mathrm{T}} \star (\Delta^2 S_n U, \Delta^2 S_{n+1} U, \ldots, \Delta^2 S_{n+p-1} U)^{-1} \Delta S_n U.$$

We also have $\Delta S_{n+1} - B\Delta S_n = \| \Delta S_n \|_F\, \Theta_n$, where Θ_n converges to 0.

From Lemma 1.8, the ratio $\| \Delta S_{n+1} \|_F / \| \Delta S_n \|_F$ is bounded, and therefore, for all $i > 1$ $\Delta S_{n+i} = B^i \Delta S_n + \| \Delta S_n \|_F\, \Theta_n^i$ and $\Delta^2 S_{n+i} = B^i(B - I)\Delta S_n + \| \Delta S_n \|_F\, \Lambda_n^i$, where

$$\Theta_n^i = \sum_{j=0}^{i-1} \frac{\| \Delta S_{n+j} \|_F}{\| \Delta S_n \|_F} B^{i-1-j} \Theta_{n+j}$$

and

$$\Lambda_n^i = \frac{\| \Delta S_{n+i} \|_F}{\| \Delta S_n \|_F} \Theta_{n+i} + \sum_{j=0}^{i-1} \frac{\| \Delta S_{n+j} \|_F}{\| \Delta S_n \|_F} B^{i-1-j}(B - I)\Theta_{n+j}.$$

From these relations we have

$$(\Delta S_n, \Delta S_{n+1}, \ldots, \Delta S_{n+p-1}) = (\Delta S_n, B\Delta S_n, \ldots, B^{p-1}\Delta S_n) + \| \Delta S_n \|_F\, (0, \Theta_n^1, \ldots, \Theta_n^{p-1}) \tag{26}$$

and

$$(\Delta^2 S_n U, \Delta^2 S_{n+1} U, \ldots, \Delta^2 S_{n+p-1} U) = (B - I)A_n + \| \Delta S_n \|_F\, (0, \Lambda_n^1 U, \ldots, \Lambda_n^{p-1} U). \tag{27}$$

By assumptions and from Lemma 2.2, with $X = (B - I)D_n$ and $Y = (B - I)D_n + (0, \Lambda_n^1 U, \ldots, \Lambda_n^{p-1} U)$. We may prove that

$$(\Delta^2 S_n U, \Delta^2 S_{n+1} U, \ldots, \Delta^2 S_{n+p-1} U)^{-1} = ((B - I)A_n)^{-1} + \frac{1}{\| \Delta S_n \|_F} \Gamma_n \tag{28}$$

where (Γ_n) is a matrix sequence converging to 0.

Also, from (26), there exists a matrix sequence (Ω_n) converging to 0 such that

$$(\Delta S_n, \Delta S_{n+1}, \ldots, \Delta S_{n+p-1})^{\mathrm{T}} \star ((B - I)A_n)^{-1} V_n = \left((I, B, \ldots, B^{p-1})^{\mathrm{T}} \star A_n^{-1}(B - I)^{-1} V_n\right) \Delta S_n + \| \Delta S_n \|_F\, \Omega_n.$$

Since the sequence $(S_n) \in L_B$, from Theorem 1.9, there exists a matrix sequence (α_n) converging to 0 such that $S_{n+1} - S = B(S_n - S) + \| \Delta S_n \|_F\, \alpha_n$.

So $\Delta S_n = (B - I)(S_n - S) + \| \Delta S_n \|_F\, \alpha_n$, and we immediately obtain

$$H_n - S = \left(I - \left((I, B, \ldots, B^{p-1})^{\mathrm{T}} \star A_n^{-1}(B - I)^{-1} V_n\right)(B - I)\right)(S_n - S) + \| \Delta S_n \|_F\, \delta_n$$

where (δ_n) is a matrix sequence converging to 0.

Applying Lemma 2.1 it follows that $H_n - S = \| \Delta S_n \|_F \, \delta_n$, and from Theorem 1.4 the ratio $\| \Delta S_n \|_F / \| E_n \|_F$ is bounded, and hence

$$\lim_{n\to\infty} \frac{\| H_n - S \|_F}{\| S_n - S \|_F} = 0. \qquad \square$$

Using this theorem we can prove the following result

Corollary 2.4. *Let $(S_n) \in L_B$ with B nonsingular and $\| B \|_F < 1$. If there exists $c > 0$ and $N > 0$ such that $| \det(A_n) | > c \| \Delta S_n \|_F^p$, for $n > N$ then,*

$$\lim_{n\to\infty} \frac{\| H_n - S \|_F}{\| S_n - S \|_F} = 0.$$

Proof. By assumptions the matrix D_n is nonsingular and $| \det(D_n) | > c$ for all $n > N$.

Therefore, from Lemma 1.13 and by the relation $\| D_n^{-1} \|_F \leqslant p^{1/2} \| D_n^{-1} \|_2$, we can prove that $\| D_n^{-1} \|_F$ is bounded. Hence

$$\lim_{n\to\infty} \| D_n^{-1} \|_F \| \Theta_{n+i} \|_F = 0 \quad \text{for } i = 0, 1, \ldots, p-1.$$

Finally the result follows from Theorem 2.3. $\square$

From the proof of Theorem 2.3 we have the following remark:

Remark 2.5. If (S_n) is a matrix sequence converging to S such that $E_{n+1} = BE_n$ with B nonsingular and $\| B \|_F < 1$, then $\Delta S_{n+1} = B \Delta S_n$. Moreover, if the matrix A_n is nonsingular for all $n > N$ then, from (26) and (27), and by a similar proof to that of Theorem 2.3, we can prove that $H_n = S$ for $n > N$.

Example 2.6. Let us consider a nonsingular matrix $D \in M_{p\times p}(\mathbb{C})$ such that $\| D \| < 1$. Then the matrix $I - D$ is nonsingular and $(I - D)^{-1} = \sum_{i=0}^{\infty} D^i$.

Now, we consider the matrix sequence defined by $S_n = \sum_{i=0}^{n} D^i$ which satisfies $\lim_{n\to\infty} S_n = (I - D)^{-1}$ and $E_{n+1} = DE_n$. Then, if the matrix $(U, DU, \ldots, D^{p-1}U)$ is nonsingular, the matrix $A_n = D^{n+1}(U, DU, \ldots, D^{p-1}U)$ is nonsingular, and from Remark 2.5 we obtain $H_n = (I - D)^{-1}$.

In this case, the transformation H allows us to calculate $(I - D)^{-1}$. We need $O(p^4)$ arithmetic operations to compute $S_n, S_{n+1}, \ldots, S_{n+p}$ and $O(p^3)$ for computing H_n.

2.3. Implementation

We shall now propose an algorithm for implementing Henrici's transformation. We shall use the H-algorithm which was first introduced in [4, 5] in order to implement the vector sequence transformation.

Let (S_n) be a matrix sequence converging to S, we consider the following ratio of determinants:

$$H_p(n) = \begin{vmatrix} S_n & \cdots & S_{n+p} \\ g_1(n) & \cdots & g_1(n+p) \\ g_2(n) & \cdots & g_2(n+p) \\ \vdots & & \vdots \\ g_p(n) & \cdots & g_p(n+p) \end{vmatrix} \div \begin{vmatrix} 1 & \cdots & 1 \\ g_1(n) & \cdots & g_1(n+p) \\ g_2(n) & \cdots & g_2(n+p) \\ \vdots & & \vdots \\ g_p(n) & \cdots & g_p(n+p) \end{vmatrix}, \tag{29}$$

where the $(g_i(n))$ are auxiliary scalar sequences.

Let us now consider the columns S_n^l of S_n and $H_p^l(n)$ of $H_p(n)$ for $l = 1, 2, \ldots, q$.

From (29) we have

$$H_p^l(n) = \begin{vmatrix} S_n^l & \cdots & S_{n+p}^l \\ g_1(n) & \cdots & g_1(n+p) \\ g_2(n) & \cdots & g_2(n+p) \\ \vdots & & \vdots \\ g_p(n) & \cdots & g_p(n+p) \end{vmatrix} \div \begin{vmatrix} 1 & \cdots & 1 \\ g_1(n) & \cdots & g_1(n+p) \\ g_2(n) & \cdots & g_2(n+p) \\ \vdots & & \vdots \\ g_p(n) & \cdots & g_p(n+p) \end{vmatrix}.$$

Note that $g_i(n)$'s are the same for all $l = 1, 2, \ldots, p$.

Using the vector version of Sylvester's identity for the numerators and the scalar one for the denominators, Sadok [13] noticed that we obtain a recursive algorithm for computing $H_p^l(n)$:

$$\begin{cases} H_0^l(n) = S_n^l,\ g_{0,i}(n) = g_i(n), & n = 0, 1, \ldots,\ i = 1, 2, \ldots, \\ H_k^l(n) = H_{k-1}^l(n) - g_{k-1,k}(n) \dfrac{\Delta H_{k-1}^l(n)}{\Delta g_{k-1,k}(n)}, & n = 0, 1, \ldots,\ k = 1, 2, \ldots, \\ g_{k,i}(n) = g_{k-1,i}(n) - g_{k-1,k}(n) \dfrac{\Delta g_{k-1,i}(n)}{\Delta g_{k-1,k}(n)}, & n = 0, 1, \ldots,\ k = 1, 2, \ldots,\ i > k \end{cases}$$

for $l = 1, 2, \ldots, q$, where Δ operates on the upper index n.

Thus we immediately obtain the H-algorithm in the matrix case for computing the $H_k(n)$'s:

$$\begin{cases} H_0(n) = S_n,\ g_{0,i}(n) = g_i(n), & n = 0, 1, \ldots,\ i = 1, 2, \ldots, \\ H_k(n) = H_{k-1}(n) - \dfrac{g_{k-1,k}(n)}{\Delta g_{k-1,k}(n)} \Delta H_{k-1}(n), & n = 0, 1, \ldots,\ k = 1, 2, \ldots, \\ g_{k,i}(n) = g_{k-1,i}(n) - \dfrac{g_{k-1,k}(n)}{\Delta g_{k-1,k}(n)} \Delta g_{k-1,i}(n), & n = 0, 1, \ldots,\ k = 1, 2, \ldots,\ i > k. \end{cases} \tag{30}$$

If we set $g_i(n) = \sum_{j=1}^{q} \Delta s_{ij}(n)u_j$ for $i = 1,2,\ldots,p$, then $H_n = H_p(n)$ is the extension of Henrici's transformation to the matrix case.

Remark 2.7. (1) For all $n \geqslant 0$, we need $O(pq)$ arithmetic operations to compute $g_{0,i}(n)$ for $i = 1,2,\ldots,p$, and $O(p^2q)$ for computing $H_n = H_p(n)$ from the terms $S_n, S_{n+1},\ldots,S_{n+p}$ of initial sequence (S_n).

(2) The algorithm (30) is well defined if

$$\begin{vmatrix} \Delta g_1(n) & \cdots & \Delta g_1(n+i-1) \\ \Delta g_2(n) & \cdots & \Delta g_2(n+i-1) \\ \vdots & & \vdots \\ \Delta g_i(n) & \cdots & \Delta g_i(n+i-1) \end{vmatrix} \neq 0, \quad \text{for } i = 1,2,\ldots,p.$$

2.4. Numerical examples

The examples of this section consist in calculating the partial sum of a power of a matrix, solving the matrix equation $X - AXD = C$, and inverting $I - D$ for every nonsingular matrix D such that $\| B \| < 1$.

We shall use the Frobenius matrix norm and the following notations:

$$R1(n) = \| S_n - S \|_F, \qquad R2(n) = \| H_n - S \|_F \quad \text{and} \quad R(n) = R2(n)/R1(n).$$

Example 2.8. Let us consider the power series of matrix defined by $S = \sum_{i=1}^{\infty} Z_i X^i$ where

$$Z_i = \begin{pmatrix} 1 & (-\frac{1}{2})^i & 1 \\ (\frac{2}{3})^i & 1 & (\frac{3}{5})^i \end{pmatrix} \quad \text{and} \quad X = \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & -\frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{3} \end{pmatrix}.$$

We have

$$S = \begin{pmatrix} 1 & \frac{1}{3} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{3} & \frac{1}{4} \end{pmatrix}$$

and, if we consider the matrix sequence (S_n) defined by $S_n = \sum_{i=1}^{n} Z_i X^i$, then we may prove that there exists an unique matrix

$$B = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \end{pmatrix}$$

such that $S_n \in L_B$ and that the hypotheses of Corollary 2.4 are satisfied.

Applying Henrici's transformation to S_n, with $U = (1,1,1)^T$ we obtain the results which are in Table 1.

Table 1

n	$R1_n$	$R2_n$	R_n
20	$0.314 \cdot 10^{-7}$	$0.545 \cdot 10^{-10}$	$0.173 \cdot 10^{-2}$
21	$0.157 \cdot 10^{-7}$	$0.180 \cdot 10^{-10}$	$0.114 \cdot 10^{-2}$
22	$0.785 \cdot 10^{-8}$	$0.799 \cdot 10^{-11}$	$0.762 \cdot 10^{-3}$
23	$0.392 \cdot 10^{-8}$	$0.198 \cdot 10^{-11}$	$0.506 \cdot 10^{-3}$
24	$0.196 \cdot 10^{-8}$	$0.660 \cdot 10^{-12}$	$0.336 \cdot 10^{-3}$
25	$0.981 \cdot 10^{-9}$	$0.219 \cdot 10^{-12}$	$0.223 \cdot 10^{-3}$
26	$0.490 \cdot 10^{-9}$	$0.729 \cdot 10^{-13}$	$0.148 \cdot 10^{-3}$
27	$0.245 \cdot 10^{-9}$	$0.241 \cdot 10^{-13}$	$0.985 \cdot 10^{-4}$
28	$0.122 \cdot 10^{-9}$	$0.791 \cdot 10^{-14}$	$0.645 \cdot 10^{-4}$

Example 2.9. Let $S = \sum_{i=1}^{\infty} Z_i X^i$ where

$$Z_i = \begin{pmatrix} i & i & i \\ 1 & (\frac{4}{3})^i & (\frac{3}{5})^i \end{pmatrix} \quad \text{and} \quad X = \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{4} & 0 \\ 0 & 0 & \frac{1}{3} \end{pmatrix}.$$

We have

$$S = \begin{pmatrix} 2 & \frac{4}{9} & \frac{3}{4} \\ 1 & \frac{1}{2} & \frac{1}{4} \end{pmatrix}.$$

In this case we may prove that the matrix B such that S_n belongs L_B is not unique, and

$$B = \begin{pmatrix} \frac{1}{2} & b \\ 0 & d \end{pmatrix}$$

with $b, d \in \mathbb{C}$.

The numerical results are given in Table 2 with $U = (1,1,1)^{\mathrm{T}}$.

Example 2.10. We consider the matrix equation

$$X - AXD = C, \tag{31}$$

where $A \in M_{p\times p}(\mathbb{C})$, $D \in M_{q\times q}(\mathbb{C})$, $C \in M_{p\times q}(\mathbb{C})$ and $X \in M_{p\times q}(\mathbb{C})$.

Let $\lambda_1, \ldots, \lambda_p$ and $\mu_1, \ldots, \mu_q$ be the eigenvalues of A and B, respectively. Wimmer [16, p.1128] showed that if $1 - \lambda_i \mu_j \neq 0$ for $i = 1, 2, \ldots, p$, $j = 1, 2, \ldots, q$ then Eq. (31) has an unique solution.

We propose the matrix version of Henrici's transformation to solve the Eq. (31).

We first construct a matrix sequence (X_n) by

$$X_{n+1} = C + AX_nD. \tag{32}$$

We have the following theorem:

Table 2

n	$R1_n$	$R2_n$	R_n
20	$0.805 \cdot 10^{-6}$	$0.475 \cdot 10^{-8}$	$0.590 \cdot 10^{-3}$
21	$0.417 \cdot 10^{-6}$	$0.167 \cdot 10^{-9}$	$0.400 \cdot 10^{-3}$
22	$0.216 \cdot 10^{-6}$	$0.586 \cdot 10^{-9}$	$0.271 \cdot 10^{-3}$
23	$0.111 \cdot 10^{-6}$	$0.205 \cdot 10^{-10}$	$0.183 \cdot 10^{-3}$
24	$0.577 \cdot 10^{-7}$	$0.716 \cdot 10^{-10}$	$0.124 \cdot 10^{-3}$
25	$0.298 \cdot 10^{-7}$	$0.249 \cdot 10^{-11}$	$0.837 \cdot 10^{-4}$
26	$0.153 \cdot 10^{-7}$	$0.870 \cdot 10^{-11}$	$0.565 \cdot 10^{-4}$
27	$0.791 \cdot 10^{-8}$	$0.301 \cdot 10^{-12}$	$0.380 \cdot 10^{-4}$
28	$0.407 \cdot 10^{-8}$	$0.105 \cdot 10^{-12}$	$0.258 \cdot 10^{-4}$

Theorem 2.11. *If $1-\lambda_i\mu_j \neq 0$ for $i = 1,2,\ldots,p$, $j = 1,2,\ldots,q$ and if X is the solution of equation* (31) *then*

$$\| X_{n+1} - X \| \leqslant \| A \| \, \| X_n - X \| \, \| D \| .$$

Moreover, if the matrices A and B are nonsingular, then

$$\frac{1}{\| A^{-1} \| \, \| D^{-1} \|} \leqslant \frac{\| X_{n+1} - X \|}{\| X_n - X \|} \leqslant \| A \| \, \| D \| .$$

Proof. From (31) and (32) we obtain $X_{n+1} - X = A(X_n - X)D$. Moreover, if A and B are nonsingular then $X_n - X = A^{-1}(X_{n+1} - X)D^{-1}$, thus the result holds. □

From Theorem 2.11 we have the following result:

Corollary 2.12. *Suppose that $1 - \lambda_i\mu_j \neq 0$ for $i = 1,2,\ldots,p$, $j = 1,2,\ldots,q$, and let X be the unique solution of Eq.* (31).

(i) *If $\| A \| \, \| D \| < 1$ then the sequence (X_n) converges to X.*

(ii) *Moreover, if the matrices A and D are nonsingular and if the limit*

$$\lim_{n\to\infty} \frac{\| X_{n+1} - X \|}{\| X_n - X \|} = R \quad \textit{exists}$$

then

$$0 < \frac{1}{\| A^{-1} \| \, \| D^{-1} \|} \leqslant R \leqslant \| A \| \, \| D \| < 1.$$

Example 2.13. Let us consider the following example

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}, \quad D = \begin{pmatrix} \frac{1}{10} & \frac{1}{20} & \frac{1}{10} \\ \frac{1}{40} & \frac{1}{10} & \frac{1}{20} \\ \frac{1}{20} & \frac{1}{10} & \frac{1}{10} \end{pmatrix} \quad \text{and} \quad C = \begin{pmatrix} -\frac{17}{40} & -\frac{2}{5} & \frac{3}{4} \\ \frac{49}{40} & \frac{11}{20} & \frac{7}{4} \end{pmatrix}.$$

Table 3

n	$R1_n$	$R2_n$	R_n
1	$0.284 \cdot 10^1$	$0.162 \cdot 10^1$	0.570
2	$0.225 \cdot 10^1$	0.137	$0.608 \cdot 10^{-1}$
3	$0.179 \cdot 10^1$	$0.128 \cdot 10^{-1}$	$0.713 \cdot 10^{-2}$
4	$0.143 \cdot 10^1$	$0.155 \cdot 10^{-2}$	$0.108 \cdot 10^{-2}$
5	$0.113 \cdot 10^1$	$0.265 \cdot 10^{-3}$	$0.233 \cdot 10^{-3}$
6	0.905	$0.546 \cdot 10^{-4}$	$0.603 \cdot 10^{-4}$
7	0.720	$0.119 \cdot 10^{-4}$	$0.165 \cdot 10^{-4}$
8	0.573	$0.264 \cdot 10^{-5}$	$0.461 \cdot 10^{-5}$
9	0.456	$0.589 \cdot 10^{-6}$	$0.129 \cdot 10^{-5}$
10	0.362	$0.131 \cdot 10^{-6}$	$0.361 \cdot 10^{-6}$

We have $\| A \|_F \, \| D \|_F \approx 0.93$ and the solution of equation (31) is

$$X = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}.$$

Applying Henrici's method to (X_n) with

$$X_0 = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & 0 & \frac{1}{5} \end{pmatrix},$$

we obtain the results shown in Table 3.

Example 2.14. Let us now consider another example. Let

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 2 & 1 \\ 1 & 3 & 2 \end{pmatrix}, \quad D = \begin{pmatrix} \frac{1}{35} & \frac{3}{70} & \frac{1}{70} & \frac{1}{70} \\ \frac{3}{70} & \frac{1}{35} & \frac{1}{70} & \frac{1}{35} \\ \frac{1}{70} & \frac{1}{70} & \frac{1}{35} & \frac{1}{70} \\ \frac{2}{35} & \frac{1}{70} & \frac{1}{35} & \frac{3}{70} \end{pmatrix} \quad \text{and } C = \begin{pmatrix} -\frac{17}{35} & \frac{31}{35} & \frac{8}{35} & 0 \\ -\frac{39}{70} & \frac{27}{14} & \frac{13}{70} & \frac{13}{14} \\ \frac{103}{70} & \frac{13}{14} & \frac{3}{14} & -\frac{3}{70} \end{pmatrix}.$$

We have $\| A \|_F \, \| D \|_F \approx 0.84$ and the solution of Eq. (31) is

$$X = \begin{pmatrix} 1 & 2 & 1 & 1 \\ 1 & 3 & 1 & 2 \\ 3 & 2 & 1 & 1 \end{pmatrix}.$$

Applying Henrici's method to (X_n) with

$$X_0 = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{5} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & 0 & \frac{1}{5} \end{pmatrix},$$

Table 4

n	$R1_n$	$R2_n$	R_n
1	0.336	0.120	0.357
2	0.231	$0.169 \cdot 10^{-1}$	$0.731 \cdot 10^{-1}$
3	0.158	$0.929 \cdot 10^{-3}$	$0.585 \cdot 10^{-2}$
4	0.109	$0.128 \cdot 10^{-3}$	$0.118 \cdot 10^{-2}$
5	$0.749 \cdot 10^{-1}$	$0.181 \cdot 10^{-4}$	$0.242 \cdot 10^{-3}$
6	$0.514 \cdot 10^{-1}$	$0.246 \cdot 10^{-5}$	$0.478 \cdot 10^{-4}$
7	$0.353 \cdot 10^{-1}$	$0.333 \cdot 10^{-6}$	$0.943 \cdot 10^{-5}$
8	$0.242 \cdot 10^{-1}$	$0.453 \cdot 10^{-7}$	$0.186 \cdot 10^{-5}$
9	$0.166 \cdot 10^{-1}$	$0.617 \cdot 10^{-8}$	$0.369 \cdot 10^{-6}$
10	$0.114 \cdot 10^{-1}$	$0.839 \cdot 10^{-9}$	$0.732 \cdot 10^{-7}$

we obtain the results shown in Table 4.

Remark 2.15. (1) These numerical examples show that the matrix version of Henrici's method is best for accelerating the convergence of some matrix sequences:

As indicated in Remark 2.5, Example 2.6 shows that this method is exact for inverting $I - D$, and from Tables 1–4 Examples 2.8, 2.9, 2.13 and 2.14 show that the sequence H_n converges faster than the initial sequence S_n.

(2) The computation of (H_n) involves $S_n, S_{n+1}, \ldots, S_{n+p}$. Therefore, if we do not know the terms of the initial sequence (S_n), it is necessary to construct them. This construction increases the operation count of the algorithm.

(3) Before the fact that this method accelerates the convergence of some matrix sequences, one has to notice the simplicity of the algorithm (30).

Elaboration of new transformations of matrix sequences, in order to resolve some matrix equations, are under consideration and will be compared with Henrici's method.

3. Acknowledgements

I am grateful to S.Belmehdi for many helpful discussions.

References

[1] M.D. Benchiboun, Linear convergence for vector sequences and some applications, *J. Comput. Appl. Math.* **55** (1994) 81–97

[2] C. Brezinski, Accélération de la convergence en analyse numérique, Lecture Notes in Math. **584** (Springer, Berlin, 1977).

[3] C. Brezinski, Other manifestations of the Schur complement, *Linear Algebra Appl.* **111** (1988) 231–247.

[4] C. Brezinski, About Henrici's method for non linear equations, talk given at *Symposium on Numerical Analysis and Computational Complex Analysis*, Zurich, 15–17 August, 1983, unpublished.

[5] C. Brezinski and H. Sadok, Vector sequence transformation and fixed point methods, in: C. Taylor et al., Eds, *Numerical Methods in Laminar and Turbulent Flows* (Pineridje, Swansea, 1987).

[6] R.C. Busby and W. Fair, Convergence of 'periodic in the limit' operator continued fractions, *SIAM J. Math. Anal.* **10** (1979) 512–522.
[7] J.P. Delahaye, *Sequence Transformations* (Springer, Berlin, 1988).
[8] G.H. Golub and C.F. Van Loan, *Matrix Computation*, (Johns Hopkins University Press, Baltimore and London, 2nd ed., 1993) 539–560.
[9] P. Henrici, *Elements of Numerical Analysis* (Wiley, New York, 1964).
[10] Y. Nievergelt, Aitken's and Steffensen's accelerations in several variables, *Numer. Math.* **59** (1991) 295–310.
[11] N. Osada, Acceleration methods for vector sequences, *J. Comput. Appl. Math.* **38** (1991) 361–371.
[12] A.M. Ostrowski, *Solutions of Equations in Euclidean and Banach Spaces* (Academic Press, New York, 1973).
[13] H. Sadok, About Henrici's transformation for accelerating vector sequences, *J. Comput. Appl. Math.* (1990) 101–110.
[14] R.C.E. Tan and A.L. Andrew, Cumputing derivatives of eigenvalues and eigenvectors by simultaneous iteration, *IMA J. Numer. Anal.* **9** (1989) 111–122.
[15] J.H. Wilkinson, *Rounding Errors in Algebraic Processes* (Prentice-Hall, Englewood Cliffs, NJ, 1963).
[16] H.K. Wimmer, Explicit solution of the matrix equation $\sum A^i X B^i = C$, *SIAM J. Matrix Anal. Appl.* **13** (1992) 1123–1130.
[17] J. Wimp, *Sequence Transformations and Their Applications* (Academic Press, New York, 1981).
[18] P. Wynn, Acceleration techniques for iterated vector and matrix problems, *Math. Comp.* **16** (1962) 301–322.